JOURNAL OF VIROLOGY, Mar. 1989, p. 1181–1187
0022-538X/89/031181-07$02.00/0

Vol. 63, No. 3

Mutational Analysis of the Conserved Basic Domain of Human Immunodeficiency Virus *tat* Protein

JOACHIM HAUBER, MICHAEL H. MALIM, AND BRYAN R. CULLEN*

Howard Hughes Medical Institute and Department of Medicine, Box 3025, Duke University Medical Center, Durham, North Carolina 27710

Received 8 September 1988/Accepted 16 November 1988

The *tat trans*-activators encoded by the known strains of primate immunodeficiency virus share a conserved, highly basic protein domain. Mutagenesis of this sequence in the *tat* gene of human immunodeficiency virus type 1 is shown here to reduce, but not eliminate, the *trans*-activation of human immunodeficiency virus type 1-specific gene expression. The degree of inhibition is shown to vary in a dose-dependent manner and is most marked at low levels of *tat* expression. Multiple mutations of the basic domain of *tat* were found to impair both the in vivo stability and the nuclear localization of the *tat* protein. It is proposed that this protein domain serves to efficiently target the *tat* gene product to its appropriate site or substrate within the nucleus of expressing cells.

The pathogenic human retrovirus human immunodeficiency virus type 1 (HIV-1) encodes a nonstructural protein, termed *tat*, which is able to activate viral gene expression when present in *trans* (2, 5, 26, 34, 36). Functional expression of the *tat* gene is required for HIV-1 replication in vitro (7, 9). The *tat* protein is localized within the nucleus of expressing cells and acts, at least in part, by enhancing the rate of HIV-1 long terminal repeat (LTR)-specific transcription (17, 20). This transcriptional activation is in turn mediated by a sequence, termed the *trans*-activation response element, which has been mapped to the HIV-1 LTR (16, 29).

The *tat* gene of the HXB-3 strain of HIV-1 consists of two coding exons which together define a small protein of 86 amino acids (2). The N-terminal 72 amino acids of the *tat* protein, encoded by the first exon, appear sufficient for full *trans*-activation of HIV-1 LTR-specific gene expression (5, 34). Sequence comparisons between isolates of HIV-1 and the more distantly related primate immunodeficiency viruses HIV-2 and simian immunodeficiency virus (11, 14, 18) reveal two highly conserved protein sequence elements within the first exon of *tat*. The first is a 15-amino-acid stretch, extending from amino acids 22 to 37 within HXB-3, which contains several conserved cysteine residues. The second element, extending from amino acids 49 to 57 in HXB-3, consists of a highly charged, basic domain. Mutational analysis of the cysteine-rich element has demonstrated that the targeted replacement of the individual cysteines is, in all but one case, sufficient to completely ablate *tat*-mediated *trans*-activation (12, 30, 31). This result is consistent with data suggesting an essential role for these cysteine moieties in the binding of metal ions by the *tat* protein (10). In contrast, the targeted replacement of several of the conserved basic amino acids by uncharged or acidic amino acids has been reported to have little or no effect on the biological activity of *tat* in transfected cells (12, 31).

In this report, we present data demonstrating that mutations of the basic domain can significantly impair both the biological activity and the in vivo stability of the HIV-1 *tat* protein. It is proposed that this inhibition is due, at least in part, to interference with the targeting of the *tat* protein to its appropriate subcellular location.

* Corresponding author.

MATERIALS AND METHODS

Construction of molecular clones. Oligonucleotide-directed mutagenesis (35), using a bacteriophage M13 mutagenesis system (Amersham Corp., Arlington Heights, Ill.), was used to introduce targeted amino acid substitutions into the full-length, wild-type *tat* gene encoded by the expression vector pgTAT (23) (Fig. 1). The mutations introduced into pΔK and pΔR each modified four contiguous nucleotides. In the case of pΔK, the mutation generated a novel *Asp*718 restriction enzyme site (agg · aag · aag → agg · TAC · Cag), while the mutation introduced into pΔR introduced a novel *Pvu*II site (cag · cga · cga → cag · cTG · ACa). Mutant pΔKR was obtained by cleaving pΔK with *Asp*718, blunt-ending the resultant linear plasmid with Klenow DNA polymerase I, and then cleaving the DNA at a 5′ *Sal*I site. Ligation of the resultant *tat* DNA fragment with a vector DNA fragment obtained by *Sal*I and *Pvu*II cleavage of pΔR resulted in the construction of a *tat* gene lacking the sequences which encode amino acids 51 to 54 (Fig. 1). The sequences of the mutant *tat* gene expression vectors were confirmed by DNA sequence analysis. The secreted alkaline phosphatase (SEAP) indicator gene expression vector pBC12/HIV/SEAP contains a gene encoding a secreted form of human placental alkaline phosphatase under the control of a fully *tat*-responsive HIV-1 LTR, as previously described (3).

Cell culture and transfection. COS cells were maintained as previously described and were transfected by using DEAE-dextran and chloroquine (5, 6).

Assay of SEAP indicator gene expression levels. The SEAP assays were performed as previously described (3) at 60 h posttransfection. Briefly, 200-μl portions of supernatant medium were removed from the transfected cultures, heated at 65°C for 5 min, and clarified by centrifugation in a microfuge at 14,000 × *g* for 2 min. A 100-μl portion of the sampled medium was then added to 100 μl of 2× SEAP assay buffer (2 M diethanolamine [pH 9.8], 1 mM $MgCl_2$, 20 mM L-homoarginine) in a 96-well flat-bottom culture plate (Falcon) and preincubated at 37°C for 10 min. A 20-μl sample of prewarmed 120 mM *p*-nitrophenyl phosphate (Sigma Chemical Co., St. Louis, Mo.) in 1X SEAP assay buffer was then added, and the reaction was incubated with mixing at 37°C. The rate of increase in A_{405} was measured with a V_{max}

	48	49	50	51	52	53	54	55	56	57	58
HIV-1 WT	Gly	Arg	Lys	Lys	Arg	Arg	Gln	Arg	Arg	Arg	Pro
pΔK	-	-	Tyr	Gln	-	-	-	-	-	-	-
pΔR	-	-	-	-	-	-	-	Leu	Thr	-	-
pΔKR	-	-	Tyr	Δ	Δ	Δ	Δ	Leu	Thr	-	-

FIG. 1. Predicted amino acid sequence of *tat* mutants. The wild-type (WT) *tat* protein sequence is derived from the replication-competent HXB-3 strain of HIV-1 used in the parental vector pgTAT (23). The mutations were introduced by oligonucleotide-directed mutagenesis, as described in the text, and were confirmed by sequence analysis. Symbols: −, conserved amino acid; Δ, deleted amino acid.

Kinetic Microplate reader (Molecular Devices Co., Palo Alto, Calif.).

Immunoprecipitation analysis and subcellular fractionation. At 60 h posttransfection, cultures were washed with cysteine-free Dulbecco modified Eagle medium and then incubated in cysteine-free Dulbecco modified Eagle medium containing 10% dialyzed fetal calf serum for 1 h. The cells were then pulse-labeled with [^{35}S]cysteine (300 μCi/ml in cysteine-free Dulbecco modified Eagle medium) for 2 h. In some experiments, this [^{35}S]cysteine pulse was followed by a chase with excess unlabeled cysteine. Labeled cultures were harvested by using radioimmunoprecipitation assay (RIPA) buffer (6) and were then subjected to immunoprecipitation analysis with a rabbit polyclonal antipeptide antiserum directed against the first 61 coding amino acids of *tat*, as previously described (6, 23). Precipitated proteins were resolved on a discontinuous 14% sodium dodecyl sulfate-acrylamide gel and visualized by autoradiography.

Subcellular fractionation of transfected, [^{35}S]cysteine-labeled COS cells was performed essentially as described by Slamon et al. (33). At 60 h posttransfection, 60-mm COS cell culture dishes were washed three times with cold phosphate-buffered saline, scraped into phosphate-buffered saline, and pelleted by centrifugation at 500 × *g* for 10 min at 4°C. The cell pellet was gently suspended in 1.3 ml of ice-cold lysis buffer (5 mM sodium phosphate [pH 7.4], 50 mM NaCl, 150 mM sucrose, 5 mM KCl, 2 mm dithiothreitol, 1 mM $MgCl_2$, 0.5 mM $CaCl_2$, 0.1 mM phenylmethylsulfonyl fluoride, 0.1% Tergitol 15-S-9 [Union Carbide Corp., New York, N.Y.]). The resultant suspension was spun through a 300-μl cushion (consisting of 30% sucrose, 2.5 mM Tris hydrochloride [pH 7.4], 10 mM NaCl) at 1,000 × *g* to pellet the nuclei. The supernatant solution was removed and spun again in a microfuge at 14,000 × *g* for 5 min. The supernatant cytoplasmic fraction was then removed and adjusted to a 1× RIPA buffer concentration.

The nuclear pellet was washed three times (with cold 20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.5], 5 mM KCl, 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, 0.2 M sucrose) and was then lysed by incubation in RIPA buffer at ambient temperature for 15 min. The nuclear lysate was then cleared by centrifugation at 14,000 × *g* for 15 min. Samples of the cytoplasmic and nuclear lysate fractions were subjected to trichloroacetic acid precipitation (6) to determine incorporated radioactivity, and equal amounts of labeled protein were then used for immunoprecipitation with the anti-*tat* antibody, as described above.

Indirect immunofluorescence. The indirect immunofluorescence analysis of the transfected COS cultures was performed as previously described (6, 17). The primary rabbit anti-*tat* antibody was used at a 1:800 dilution, while the second antibody (rhodamine-conjugated goat anti-rabbit immunoglobulin G; Boehringer-Mannheim Biochemicals, Indianapolis, Ind.) was used at a 1:50 dilution.

RESULTS

The conserved basic domain of *tat* is required for full biological activity. To address the role of the conserved basic domain of the *tat* protein in the *trans*-activation of HIV-1 LTR-specific gene expression, we used oligonucleotide-directed mutagenesis (35) to alter this region in a wild-type *tat* expression vector, pgTAT (Fig. 1). In the pΔK mutant, lysine residues at amino acid positions 50 and 51 were replaced by tyrosine and glutamine, respectively. In the pΔR mutant, arginine residues at positions 55 and 56 were similarly replaced by leucine and threonine. Each of these mutations therefore replaced two of the eight charged amino acids in the basic region with uncharged amino acids. The more extensive mutation present in pΔKR included the substitution mutations at positions 50, 55, and 56 described for pΔK and pΔR and the deletion of amino acids present between positions 51 and 54 of the *tat* protein. The pΔKR mutant therefore lacks six of the eight basic residues present in the wild-type *tat* protein (Fig. 1).

The phenotypes of the *tat* mutants were examined by using a transient expression assay based on the SEAP indicator gene (3). A previously described (3) construction in which the SEAP gene is linked in *cis* to the HIV-1 LTR (pBC12/HIV/SEAP) was transfected into the HIV-1 replication-competent cell line COS (22) together with a range of concentrations of the *tat* expression plasmids (Fig. 2). The level of expression of the SEAP protein in the supernatant media of the transfected cultures was determined at 60 h posttransfection by using a simple, highly quantitative spectrophotometric assay for alkaline phosphatase (3). The observed levels of *tat*-mediated *trans*-activation for each concentration of *tat* vector used, expressed as a multiple of the basal level encoded by the HIV-1 LTR construction alone, are shown in Fig. 2. An intermediate level of transfection of the wild-type *tat* gene encoded by pgTAT was observed to yield the maximal observed *trans*-activation of ~90 fold. This is comparable to the maximal level of *tat*-mediated *trans*-activation previously observed by us and others using the COS cell line (6, 20). Introduction of higher or lower levels of pgTAT yielded a slightly but consistently lower level of indicator gene activity. The mutant *tat* vectors pΔR and pΔK displayed comparable levels of HIV-1 LTR *trans*-activation over the entire concentration range used; however, this activity increased dramatically with increasing levels of the transfected *tat* vector (Fig. 2). The net effect of the different dose responses observed with these mutant *tat* constructions was that the pΔK and pΔR clones displayed an essentially wild-type phenotype at the highest level of *tat* vector used (31-fold *trans*-activation versus 57-fold for pgTAT), and yet they showed a marked inhibition of *trans*-activation when the pBC12/HIV/SEAP vector was present at a 100-fold excess (2-fold versus 32-fold for pgTAT). The double mutant pΔKR displayed a similar dose dependence, yielding no detectable *trans*-activation at low levels of *tat* vector transfection but giving rise to a readily detectable ~five-fold *trans*-activation of the HIV-1 LTR at the highest level of *tat* vector tested (Fig. 2).

One possible explanation for the mutant phenotypes shown in Fig. 2 is that little or no *tat* mutant protein was

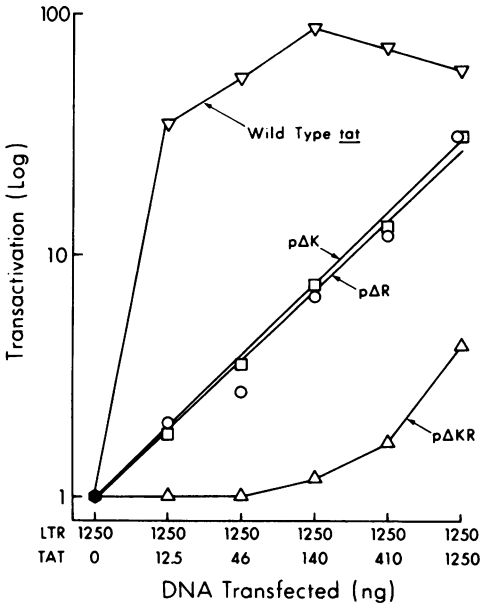

FIG. 2. Relative *trans*-activation of HIV-1 LTR-specific gene expression mediated by various levels of the mutant or wild-type *tat* gene product. A constant level (1,250 ng) of the indicator construction pBC12/HIV/SEAP (LTR) was cotransfected with increasing levels (0 to 1,250 ng) of the various *tat* gene expression plasmids (TAT) as indicated. The total amount of DNA transfected per culture was maintained at 2,500 ng by supplementation with a negative control vector, pBC12/CMV/IL-2 (5). The levels of SEAP expression in the supernatant media were determined at 60 h posttransfection, as previously described (3), and are expressed as a multiple of the basal level of HIV-1 LTR-specific SEAP gene expression.

being synthesized in the transfected cells. To test this hypothesis, we performed quantitative immunoprecipitations of [^{35}S]cysteine-labeled transfected COS cultures with a previously described rabbit polyclonal anti-*tat* antiserum (17). As previously noted, the wild-type 86-amino-acid *tat* protein migrates in sodium dodecyl sulfate-acrylamide gels at an unexpectedly high relative molecular weight (M_r) of ~15,500 (Fig. 3A) (23, 25). Interestingly, both the pΔK and pΔR mutants, which are of the same predicted size as the wild-type *tat* protein, migrate at a somewhat lower M_r of ~14,500. This suggests that the highly basic nature of the HIV-1 *tat* protein is at least partially responsible for the aberrant migration of the *tat* protein observed in sodium dodecyl sulfate-acrylamide gels. This possibility is supported by the migration behavior of the *tat* protein encoded by pΔKR, which has suffered a more extensive mutation of the conserved basic region. This protein is predicted to be only 4 amino acids shorter than the pΔK and pΔR mutants, and yet it migrates significantly more rapidly (M_rs of ~13,000 versus ~14,500). In total, the rate of synthesis of each of the *tat* mutants in the transfected cultures was comparable, although the culture transfected with pΔKR did appear to yield a slightly reduced signal (Fig. 3A).

Mutations in the conserved basic domain affect the subcellular localization of the *tat* protein. Because the basic region of the *tat* protein possesses amino acid homology to nuclear transport signals previously noted for other, predominantly nuclear proteins (see below), we next examined the subcellular localization of the mutant *tat* proteins. In this experiment, transfected cells were labeled with [^{35}S]cysteine as described in the legend to Fig. 3A and were then subjected to a 1-h chase with excess unlabeled cysteine to permit labeled proteins to fully localize to the appropriate subcellular compartment. Total cytoplasmic and nuclear proteins were then separated by using the fractionation protocol developed by Slamon et al. (33). The incorporated radioactivity in each fraction was determined, and equal amounts were then used for immunoprecipitation with anti-*tat* antiserum. The *tat* protein encoded by the parental vector pgTAT, as well as the mutant *tat* molecules encoded by the pΔK and pΔR vectors, was localized to the nucleus within 1 h of being labeled with [^{35}S]cysteine (Fig. 3B). In sharp contrast, the ~13-kilodalton *tat* protein encoded by pΔKR appeared to remain predominantly within the cytoplasm of expressing cells. This result therefore indicates that the mutation present in pΔKR must perturb a protein sequence necessary for the nuclear localization of the encoded *tat* protein.

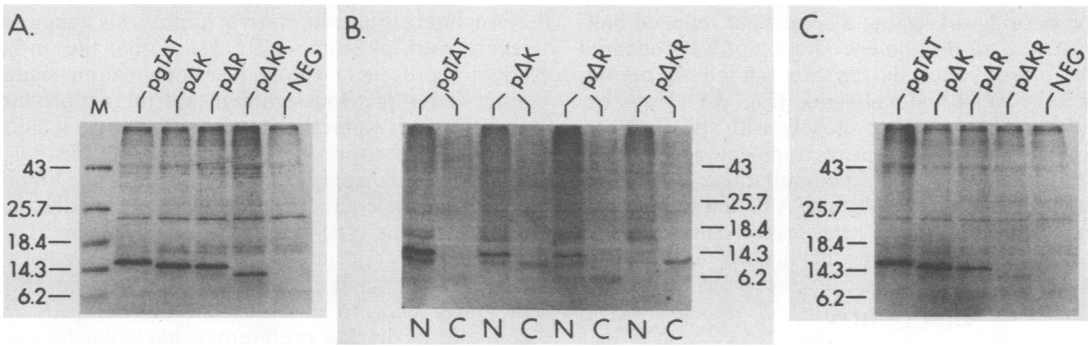

FIG. 3. Analysis of *tat* protein expression by immunoprecipitation analysis. COS cells were transfected (6) with equal levels of the various *tat* expression vectors and were then metabolically labeled with [^{35}S]cysteine for 2 h at 60 h posttransfection. Labeled *tat* proteins were isolated by immunoprecipitation with a polyclonal rabbit anti-*tat* antiserum (6), resolved on a discontinuous 14% sodium dodecyl sulfate-acrylamide gel, and visualized by autoradiography. The vector transfected into each culture is indicated over the appropriate gel lane. The negative control (NEG) used was the vector pBC12/CMV/IL-2 (5). ^{14}C-labeled protein molecular size markers (M) were obtained from Bethesda Research Laboratories, Inc., Gaithersburg, Md., and are indicated in kilodaltons. (A) Cultures were harvested directly after the [^{35}S]cysteine labeling pulse to determine the rate of *tat* protein synthesis. (B) Cultures were chased with media containing excess unlabeled cysteine for 1 h after completion of the [^{35}S]cysteine pulse. Transfected cells were then separated (33) into nuclear (N) and cytoplasmic (C) fractions before immunoprecipitation. (C) The stability of the mutant *tat* proteins was examined by exposing the transfected COS cultures to excess unlabeled cysteine for 6 h after completion of the [^{35}S]cysteine pulse before harvest and immunoprecipitation. The experiment for panel C was performed in parallel to that for panel A and was exposed to autoradiography for the same period of time.

The immunoprecipitation analysis presented in Fig. 3B examined the subcellular transport of newly synthesized *tat* protein. To extend this analysis to an examination of the subcellular localization of the various *tat* proteins at steady state, we also performed an indirect immunofluorescence analysis of transfected COS cell cultures using the rabbit anti-*tat* antiserum (17). These results confirmed our previous observation that the wild-type *tat* protein is concentrated within the nucleus of expressing cells (Fig. 4A). The *tat* protein encoded by pΔK yielded the same subcellular distribution pattern as the wild-type *tat* protein, including the previously described nucleolar concentration (17) (Fig. 4B). Neither the pgTAT- nor the pΔK-transfected cells yielded any detectable cytoplasmic fluorescence. Analysis of cells transfected with pΔR yielded a fluorescence signal which was consistently fainter than the signal observed for the wild-type pgTAT vector or the pΔK vector (Fig. 4C versus A and B). Compensation for this weaker signal by optimization of the photographic exposure time demonstrated that the *tat* protein encoded by pΔR was concentrated within the nuclei and nucleoli of expressing cells but also revealed the presence of a low level of cytoplasmic *tat* protein (Fig. 4D). In contrast, the immunofluorescence analysis of COS cell cultures transfected with the extensively mutated *tat* expression vector pΔKR yielded no detectable fluorescence signal above background level in either subcellular compartment (Fig. 4E and F).

Mutagenesis of the conserved basic region affects the stability of the *tat* protein. Our inability to detect the mutant *tat* protein encoded by pΔKR using an assay which measures the steady-state level of protein expression (Fig. 4) suggested the possibility that mutagenesis of the conserved basic domain might affect *tat* protein stability in vivo. To test this hypothesis, transfected COS cultures were labeled with [^{35}S]cysteine as described for Fig. 3A and then exposed for 6 h to media containing large excesses of unlabeled cysteine. This assay permitted the detection of significant differences in the in vivo half-lives of the various *tat* proteins. Thus, the wild-type *tat* protein encoded by pgTAT and the mutant *tat* protein encoded by pΔK showed little if any decline in signal intensity during the chase period, suggesting that these proteins are fairly stable in vivo (Fig. 3C). The pΔR mutant exhibited a slight, ~two- to three-fold drop in signal intensity over the same period, suggesting a somewhat reduced half-life. In contrast, the *tat* protein encoded by pΔKR appeared to be rapidly turned over during the 6-h chase period, resulting in a very weak residual signal (Fig. 3A versus C). These results therefore correlate closely with the data presented in Fig. 4 in that the culture transfected with pΔR presented a reduced *tat*-specific immunofluorescence signal relative to the signals produced by pgTAT- and pΔK-transfected cultures, while the pΔKR-transfected culture in that case yielded no detectable *tat*-specific signal.

DISCUSSION

The *tat trans*-activators encoded by the known primate immunodeficiency viruses display extensive amino acid sequence heterogeneity and yet demonstrate the ability to cross-*trans*-activate virus-specific gene expression (1, 11, 14). This observed sequence heterogeneity may therefore be viewed as an experiment by nature and focuses attention on those sequence domains which have been conserved across the known primate virus isolates. The first of these, a highly conserved cysteine-rich domain, has been proposed to be important in the binding of metal ions by *tat* and in a subsequent protein dimerization event (10). The targeted mutagenesis of this domain was recently shown to result in the loss of all *trans*-activation activity by *tat* (12, 30, 31). In this study, we have examined the functional consequences of mutations within the second conserved *tat* domain, a highly basic region centered at approximately position 53 in the *tat* amino acid protein sequence. We demonstrate that these sequence changes result in an inhibition of *tat*-mediated *trans*-activation which is partially alleviated by high levels of *tat* expression. We also demonstrate that mutagenesis of this domain can result in two detectable biochemical lesions, i.e., interference with the nuclear localization of *tat* and the destabilization of the *tat* protein.

The rapid accumulation of the wild-type *tat* protein within the nucleus of expressing cells (Fig. 3B) suggests the presence of a nuclear localization signal in the primary sequence of this protein. Ruben et al. (30) have in fact recently demonstrated that a 5-amino-acid sequence derived from the basic domain of *tat* (i.e., amino acids 48 to 52) is sufficient to induce the nuclear localization of the normally cytoplasmic β-galactosidase gene product when introduced at the amino terminus. A comparison of the conserved basic domain with the sequences of known nuclear localization signals reveals a high degree of homology in several instances (Fig. 5A). Of interest is the observation that this homology appears to extend over only part of the basic domain in each case. Thus, the well-defined nuclear localization signal of the simian virus 40 large T antigen (19) displays a close homology to amino acids 48 to 53 of *tat*, while the homology between one of the polyomavirus large T antigen nuclear localization signals (28) and *tat* is confined to amino acids 54 to 60. One interpretation of these sequence comparisons, which would be fully consistent with the data presented in this report, is that the basic region actually consists of two adjacent or partially overlapping nuclear localization signals. Mutagenesis of either part alone, i.e., pΔK or pΔR, would therefore have little effect on the final nuclear localization of *tat*, and only a double mutant, i.e., pΔKR, would then yield the predicted nonnuclear phenotype. Multiple nuclear localization signals have been observed in several nuclear proteins which have been examined in detail (13, 15, 27, 28) and this redundancy has been proposed to enhance the rate of nuclear protein transport (21).

An interesting alternative hypothesis, suggested by the recent work of Siomi et al. (32), is that the *tat* basic region might represent a subnuclear localization signal. We and others have previously noted that the *tat* protein is preferentially, and sometimes predominantly, localized to the nucleolus of expressing cells (17, 30). Further, the basic domain of *tat* displays remarkable sequence homology to the proposed nucleolar localization signal of the human T-cell leukemia virus type 1 *rex* gene product (Fig. 5B), which has also been proposed to consist of two adjacent nuclear localization signals (32). However, the observation that both the pΔK and, to a lesser extent, the pΔR *tat* proteins continue to display preferential nucleolar localization (Fig. 4) may argue against this attractive hypothesis.

A second effect of mutations within the *tat* basic domain, observed to a slight degree in pΔR and to a marked degree in pΔKR, is a reduction in the in vivo stability of the encoded protein. It is of interest that the stabilities of the various encoded *tat* proteins in fact appear to correlate with their subcellular localization as derived from data shown in Fig. 3B and 4. Thus, the predominantly nuclear pΔK mutant appears to share the long half-life of the parental *tat* protein, while the predominantly cytoplasmic *tat* protein encoded by

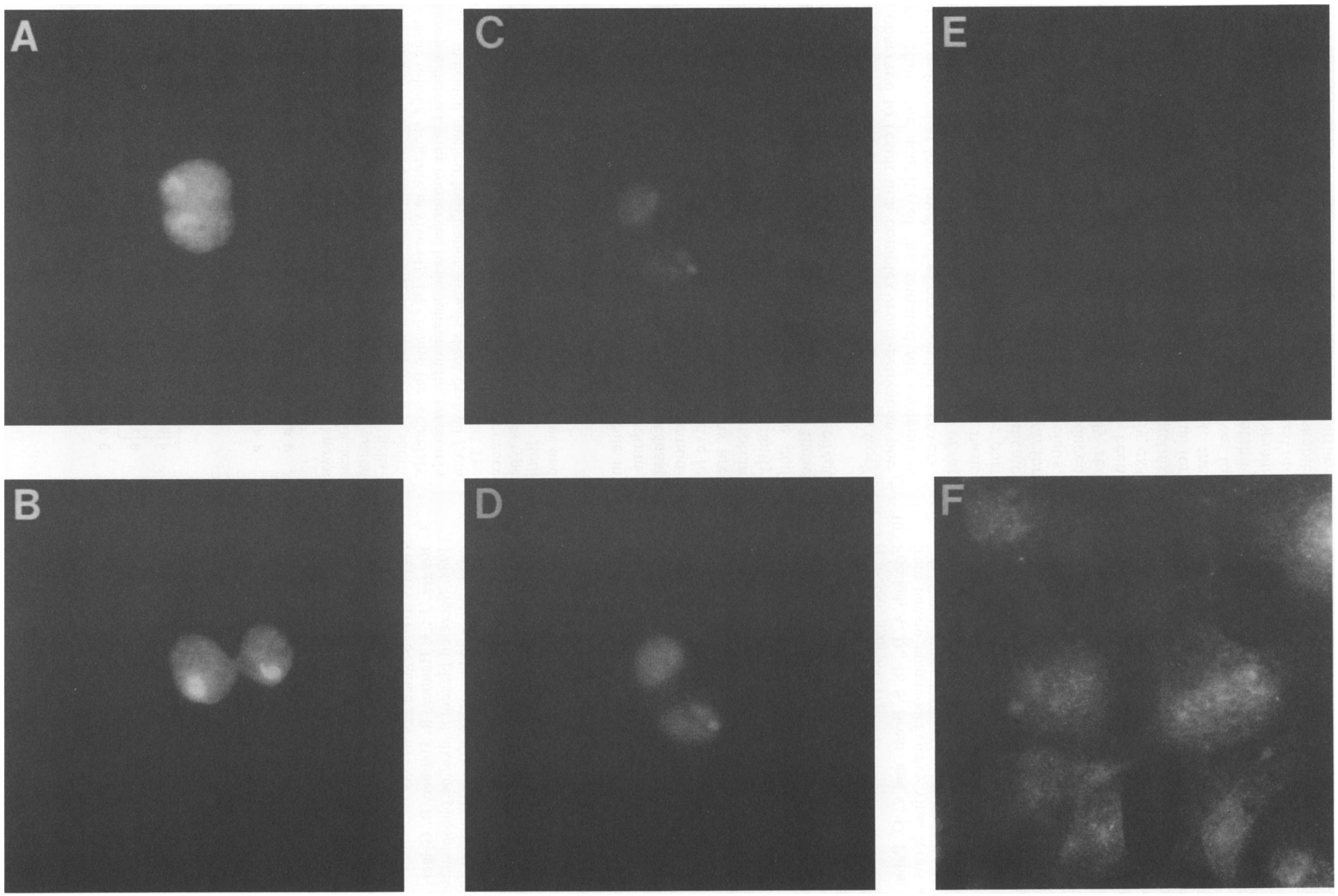

FIG. 4. Subcellular localization of wild-type and mutant *tat* proteins visualized by indirect immunofluorescence. Wild-type *tat* protein expression was detected in ~5% of the transfected COS cells, as expected (16). The photographs shown were judged representative of transfected cultures as a whole. (A, B, C and E) Photographs taken with the same exposure time (0.32 s). Longer exposures were used to more clearly reveal the subcellular distribution of pΔR (D) (1.25 s) or to visualize cells present in a negative field (F) (3.46 s). (A) pgTAT (wild-type);(B) pΔK; (C and D) pΔR; (E and F) pΔKR.

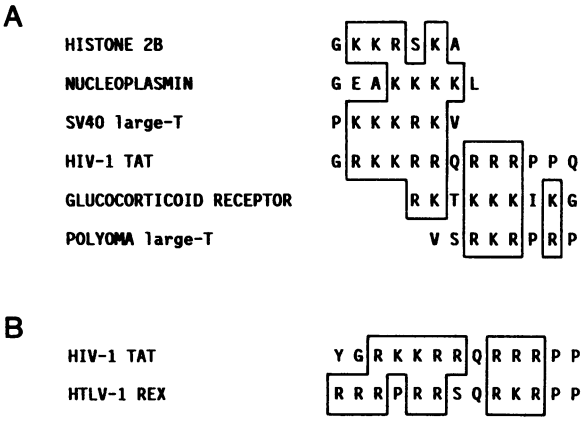

FIG. 5. Homology of the conserved basic region of *tat* to known nuclear (A) and nucleolar (B) localization signals. (A) Nuclear localization signals in general contain a stretch of the basic amino acids lysine and arginine (boxed) flanked by one or more helix-breaking amino acids (proline or glycine) and may form a short, highly basic α-helix structure in vivo (4). This is also the secondary structure predicted for the conserved basic domain of the *tat* protein (2). (The indicated nuclear localization signals were derived from references 4, 19, 24, 27, and 28). SV40, Simian virus 40. (B) The basic region of the HIV-1 *tat* protein displays extensive sequence homology to the recently defined human T-cell leukemia virus type 1 (HTLV-1) *rex* protein nucleolar localization signal (32). Both contain eight basic amino acids (boxed) and contain glutamine and proline residues at equivalent positions.

the pΔKR mutant appears to be significantly more labile in vivo. The *tat* protein encoded by pΔR, which is predominantly localized within the cell nucleus yet displays at least some cytoplasmic expression (Fig. 4D), in turn demonstrated a slightly but significantly reduced half-life relative to that of the wild-type *tat* protein. These observations are therefore consistent with the hypothesis that the *tat* protein is stable when localized to the cell nucleus but significantly more labile when inappropriately expressed in the cell cytoplasm. Clearly, however, the correlation between these two *tat* mutant phenotypes does not prove a cause-and-effect relationship.

The mutations introduced into the *tat* protein in this work were observed to result in a complex *trans*-activation phenotype. Thus, the activity of the mutated *tat* proteins is markedly reduced at low levels of expression but recovers, in two cases to an almost wild-type level, at high levels of *tat* protein expression. This suggests that the conserved basic domain, unlike the cysteine domain, is not an intrinsic part of the active site of the *tat* protein and may instead indicate that this sequence functions to enhance the affinity of *tat* for its appropriate site or substrate within the cell nucleus. Thus, mutagenesis of this domain would result in a reduced *trans*-activation activity which could be significantly restored by increasing the concentration of *tat*. It has been reported that only very low levels of *tat* are actually expressed within HIV-1-infected cells (8), so that the targeting of *tat* to its appropriate cellular location may be essential for optimum biological activity in vivo. In this context, the hypothesis that the conserved basic domain of *tat* serves to enhance the nuclear transport, and hence the in vivo stability, of this protein appears only partly sufficient to explain the *trans*-activation phenotypes displayed by the mutated *tat* proteins. Thus, the pΔK mutant, which is not distinguishable from the wild type by these criteria (Fig. 3 and 4), nevertheless fails to demonstrate full *trans*-activation (Fig. 2). We therefore favor the hypothesis that the basic domain serves a second, as yet undefined, targeting function within the nucleus of the expressing cell. Clearly, a resolution of this question will require a fuller understanding of the mechanism of action of *tat* and, in particular, the reconstitution of *tat*-mediated *trans*-activation in a defined in vitro system.

ACKNOWLEDGMENTS

We thank Sharon Goodwin for excellent secretarial assistance and Craig Rosen for communication of results before publication.

LITERATURE CITED

1. **Arya, S. K., B. Beaver, L. Jagodzinski, B. Ensoli, P. J. Kanki, J. Albert, E.-M. Fenyo, G. Biberfeld, J. F. Zagury, F. Laure, M. Essex, E. Norrby, F. Wong-Staal, and R. C. Gallo.** 1987. New human and simian HIV-related retroviruses possess functional transactivator (tat) gene. Nature (London) **328**:548–550.
2. **Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal.** 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). Science **229**:69–73.
3. **Berger, J., J. Hauber, R. Hauber, R. Geiger, and B. R. Cullen.** 1988. Secreted placental alkaline phosphatase: a powerful new quantitative indicator of gene expression in eukaryotic cells. Gene **66**:1–10.
4. **Burglin, T. R., and E. M. De Robertis.** 1987. The nuclear migration signal of Xenopus laevis nucleoplasmin. EMBO J. **6**:2617–2625.
5. **Cullen, B. R.** 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell **46**:973–982.
6. **Cullen, B. R.** 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. Methods Enzymol. **152**:684–703.
7. **Dayton, A. I., J. G. Sodroski, C. A. Rosen, W. C. Goh, and W. A. Haseltine.** 1986. The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. Cell **44**:941–947.
8. **Feinberg, M. B., R. F. Jarrett, A. Aldovini, R. C. Gallo, and F. Wong-Staal.** 1986. HTLV-III expression and production involve complex regulation at the level of splicing and translation of viral RNA. Cell **46**:807–817.
9. **Fisher, A. G., M. B. Feinberg, S. F. Josephs, M. E. Harper, L. M. Marselle, G. Reyes, M. A. Gonda, A. Aldovini, C. Debouk, R. C. Gallo, and F. Wong-Staal.** 1986. The trans-activator gene of HTLV-III is essential for virus replication. Nature (London) **320**:367–371.
10. **Frankel, A. D., D. S. Bredt, and C. O. Pabo.** 1988. Tat protein from human immunodeficiency virus forms a metal-linked dimer. Science **240**:70–73.
11. **Fukasawa, M., T. Miura, A. Hasegawa, S. Morikawa, H. Tsujimoto, K. Miki, T. Kitamura, and M. Hayami.** 1988. Sequence of simian immunodeficiency virus from African green monkey, a new member of the HIV/SIV group. Nature (London) **333**: 457–461.
12. **Garcia, J. A., D. Harrich, L. Pearson, R. Mitsuyasu, and R. B. Gaynor.** 1988. Functional domains required for *tat*-induced transcriptional activation of the HIV-1 long terminal repeat. EMBO J. **7**:3143–3147.
13. **Greenspan, D., P. Palese, and M. Krystal.** 1988. Two nuclear location signals in the influenza virus NS1 nonstructural protein. J. Virol. **62**:3020–3026.
14. **Guyader, M., M. Emerman, P. Sonigo, F. Clavel, L. Montagnier, and M. Alizon.** 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. Nature (London) **326**:662–669.
15. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of E. coli β-galactosidase to the nucleus in yeast. Cell **36**:1057–1065.
16. **Hauber, J., and B. R. Cullen.** 1988. Mutational analysis of the *trans*-activation-responsive region of the human immunodeficiency virus type I long terminal repeat. J. Virol. **62**:673–679.

17. **Hauber, J., A. Perkins, E. P. Heimer, and B. R. Cullen.** 1987. Trans-activation of human immunodeficiency virus gene expression is mediated by nuclear events. Proc. Natl. Acad. Sci. USA **84**:6364–6368.
18. **Hirsch, V., N. Riedel, and J. I. Mullins.** 1987. The genome organization of STLV-3 is similar to that of the AIDS virus except for a truncated transmembrane protein. Cell **49**:307–319.
19. **Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear location. Cell **39**:499–509.
20. **Kao, S.-Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin.** 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature (London) **330**:489–493.
21. **Lanford, R. E., P. Kanda, and R. C. Kennedy.** 1986. Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. Cell **46**:575–582.
22. **Levy, J. A., C. Cheng-Mayer, D. Dina, and P. A. Luciw.** 1986. AIDS retrovirus (ARV-2) clone replicates in transfected human and animal fibroblasts. Science **232**:998–1001.
23. **Malim, M. H., J. Hauber, R. Fenrick, and B. R. Cullen.** 1988. Immunodeficiency virus *rev trans*-activator modulates the expression of the viral regulatory genes. Nature (London) **335**:181–183.
24. **Moreland, R. B., G. L. Langevin, R. H. Singer, and R. L. Garcea, and L. M. Hereford.** 1987. Amino acid sequences that determine the nuclear localization of yeast histone 2B. Mol. Cell. Biol. **7**:4048–4057.
25. **Muesing, M. A., D. H. Smith, and D. J. Capon.** 1987. Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell **48**:691–701.
26. **Peterlin, B. M., P. A. Luciw, P. J. Barr, and M. D. Walker.** 1986. Elevated levels of mRNA can account for the trans-activation of human immunodeficiency virus. Proc. Natl. Acad. Sci. USA **83**:9734–9738.
27. **Picard, D., and K. R. Yamamoto.** 1987. Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. EMBO J. **6**:3333–3340.
28. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44**: 77–85.
29. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41**:813–823.
30. **Ruben, S., A. Perkins, R. Purcell, K. Joung, R. Sia, R. Burghoff, W. A. Haseltine, and C. A. Rosen.** 1989. Structural and functional characterization of human immunodeficiency virus *tat* protein. J. Virol. **63**:1–8.
31. **Sadaie, M. R., J. Rappaport, T. Benter, S. F. Josephs, R. Willis, and F. Wong-Staal.** 1988. Missense mutations in an infectious HIV genome: functional mapping of tat and demonstration of a novel trs splice acceptor. Proc. Natl. Acad. Sci. USA **85**: 9224–9228.
32. **Siomi, H., H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka.** 1988. Sequence requirements for nucleolar localization of human T cell leukemia virus type 1 pX protein, which regulates viral RNA processing. Cell **55**:197–209.
33. **Slamon, D. J., W. J. Boyle, D. E. Keith, M. F. Press, D. W. Golde, and L. M. Soura.** 1988. Subnuclear localization of the *trans*-activating protein of the human T-cell leukemia virus type I. J. Virol. **62**:680–686.
34. **Sodroski, J., R. Patarca, and C. Rosen.** 1985. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science **229**:74–77.
35. **Taylor, J. W., J. Ott, and F. Eckstein.** 1985. The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. Nucleic Acids Res. **13**:8765–8785.
36. **Wright, C. M., B. K. Felber, H. Paskalis, and G. N. Pavlakis.** 1986. Expression and characterization of the trans-activator of HTLV-III/LAV virus. Science **234**:988–992.